*Visual Onto-thesaurus for Dravidian languages (a brief note)

Dr. Rajendran Sankaravelayuthan &

Dr. Anandkumar, M.

* The topic “Visual-thesaurus for Dravidian languages” is only a proposal we plan to implement, probably with the help of other institutions.
* At present we have developed Onto-thesaurus for Tamil which we are planning to expand as Visual Onto-thesaurus for Tamil.
* In the present talk focuses on the theoretical and practical foundation in building Onto-thesaurus for Tamil which can be easily extended for other Dravidian languages.

* Visual-thesaurus for Dravidian languages

* Tamil Visual Onto-thesaurus (TVOT or simply VOT) is an outcome of an extensive research activity that went on in the field of lexical semantics of Tamil.
* It went through several stages before being culminated into Tamil visual onto-thesaurus.
* It depicts our travel from Tamil thesaurus to Tamil word net.
* It is a lexical resource which amalgamates all sorts of information available in a dictionary, thesaurus and wordNet.
* The Dravidian wordNets (in which Tamil wordNet is one of the four components) built under the IndoWordNet project depended on an ontology developed by Western conceptualization of the world found in English).
* This has not taken into consideration the Indian conceptualization of the world depicted in the *nikhandu* tradition.

* Background of Visual onto-thesaurus

* Say for examples, *nikhandus* have classifications such a six types of tastes, nine types of planets (*gragams*), 7 types of *mandalams* (a type of division), 15 *tidis* (15 phases of moon), etc. which are crucial for Indian tradition. In the western oriented WordNet ontology there is no scope for the visualization of concepts depicted in *nikhandus*.
* Moreover building a wordNet based on Hindi wordNet which in turn is built on English wordNet will take many years to complete and it would miss the conceptualization depicted in Indian tradition. Apart from this the extension approach of building Tamil wordNet using Hindi wordNet cannot fulfill Dravidian conceptualization.
* A merger approach of building separate wordNets and collapsing them into one would have been a preferable approach. The present visual onto-thesaurus is based on the Indian and Dravidian conceptualization and the process of building one is comparatively very simple.
* We have the plan to mend it into a generic one so that all the Dravidian languages can be easily accommodated into it.

* Background of Visual onto-thesaurus

* A paper thesaurus for Tamil was prepared in 1990 based on the principles of componential analysis of meaning propounded by Nida (1975), Indian tradition of *nikhandu* and Aristotelian principle of genera and species and was published in 2001 (Rajendran, 2001), nearly after a decade.
* Following the paper thesaurus, an Electronic thesaurus for Tamil was attempted and a book on Tamil electronic thesaurus was published in 2006 (Rajendran and Baskaran, 2006).
* The preparation of wordNet for Tamil was undertaken (2001-2003) with the financial assistance from Tamil Virtual University (renamed now as Tamil virtual academy) and a crude version of it based on the ontology developed by Rajendran (Rajendran, 2001) was submitted to the institute in 2003.
* After that, from 2009 onwards with the fund received from MHRD and Department of electronics and information Technology of Govt. of India the building of Dravidian wordNets were executed based on Hindi wordNet; nearly 30000 synsets (concepts) have been completed. Still we have a long way to go to achieve the desired target.
* At present a team from CEN, Amrita University is involved in building onto-thesaurus for Tamil as a part of the project entitled “Computing Tools for Tamil Language teaching and learning”. The project is funded by Tamil Virtual Academy, Chennai.

* Background of Visual onto-thesaurus

* Ontology is the philosophical study of the nature of being, becoming, existence or reality, as well as the basic categories of being and their relations.
* Traditionally listed as a part of the major branch of philosophy known as metaphysics, ontology often deals with questions concerning what entities exist or may be said to exist and how such entities may be grouped, related within a hierarchy, and subdivided according to similarities and differences.
* Although ontology as a philosophical enterprise is highly hypothetical, it also has practical application in information science and technology, such as ontology engineering. (From Wikipedia)
* The word ontology is from the Greek *ontos* for being and *logos* for word. One classical dictionary definition of Ontology is: "The branch of metaphysics that deals with the nature of being".

* Ontology

* It originated with Aristotle’s effort to classify things in the world. Ontology is an explicit specification of a conceptualization. As a term borrowed from philosophy, ontology means a systematic account of ‘existence’.
* AI systems deliberates that what "exists" can be represented in a program code. They try to represent knowledge of a domain in a declarative formalism; the set of objects that can be represented is called the universe of discourse.

* Hyponymy and its consequence taxonomy are the fundamental building blocks of ontology.
* Hyponymy and its natural partner, incompatibility, are described by Lyons (1977) as “the most fundamental paradigmatic relations of sense it terms of which the vocabulary is structured”.
* Lyons (1977) states that taxonomic lexical hierarchies are structured by the relations of hyponymy and incompatibility.
* The relation of hyponymy imposes a hierarchical structure upon the vocabulary and upon particular fields within the vocabulary; and the hierarchical ordering of lexemes can be represented formally as a tree diagram (Lyons, 1977: 295).
* It is hard to conceive of any language operating satisfactorily in any culture without its vocabulary being structured in terms of the complementary principles of hyponymy and contrast (Lyons, 1977: 300).
* Meronymic or partonomic relations are ontological relations that are considered as fundamental as the ubiquitous, taxonomic subsumtion relationship (Pribbenow, 2002: 35).
* There are numerous lexemes in the vocabularies of languages whose meaning cannot be specified independently of some part-whole relation of sense (Lyons, 1977: 314).

* Principles of ontology

* Taxonomy is usually only a hierarchy of concepts (i.e. the only relation between the concepts is parent/child, or sub-class/super-class, or broader/narrower), but in an ontology, arbitrary complex relations between concepts can be expressed too (X married to Y; or A works for B; or C is located in D, etc.).
* Although taxonomy contributes to the semantics of a term in a vocabulary, ontologies include richer relationships between terms. It is these rich relationships that enable the expression of domain specific knowledge, without the need to include domain-specific terms.
* Ontology has a richer internal structure as it includes relations and constraints between the concepts.
* Ontology claims to represent a certain consensus about the knowledge in the domain. This consensus is among the intended users of the knowledge, e.g. doctors using a hospital ontology regarding a certain disease, artists relating to historical art and so on.
* Word vocabulary and ontology are often used interchangeably. But a more strict definition is that a vocabulary is a collection of terms being used in a particular domain that can be structured (e.g. hierarchically) as a taxonomy.

* Principles of ontology

* This taxonomy when combined with some relationships, constraints and rules, form the ontology. A combination of ontology together with a set of instances of classes constitutes a knowledge base as given below (Breslin et al, 2009: 58):
  * Vocabulary + structure = Taxonomy
  * Taxonomy + Relationships, constraints and rules = Ontology
  * Ontology + instances = Knowledge base

* **Principles of ontology**

*Aristotle instigated the history of scientific taxonomy.

*It is predicated on a first philosophy of essentialism.

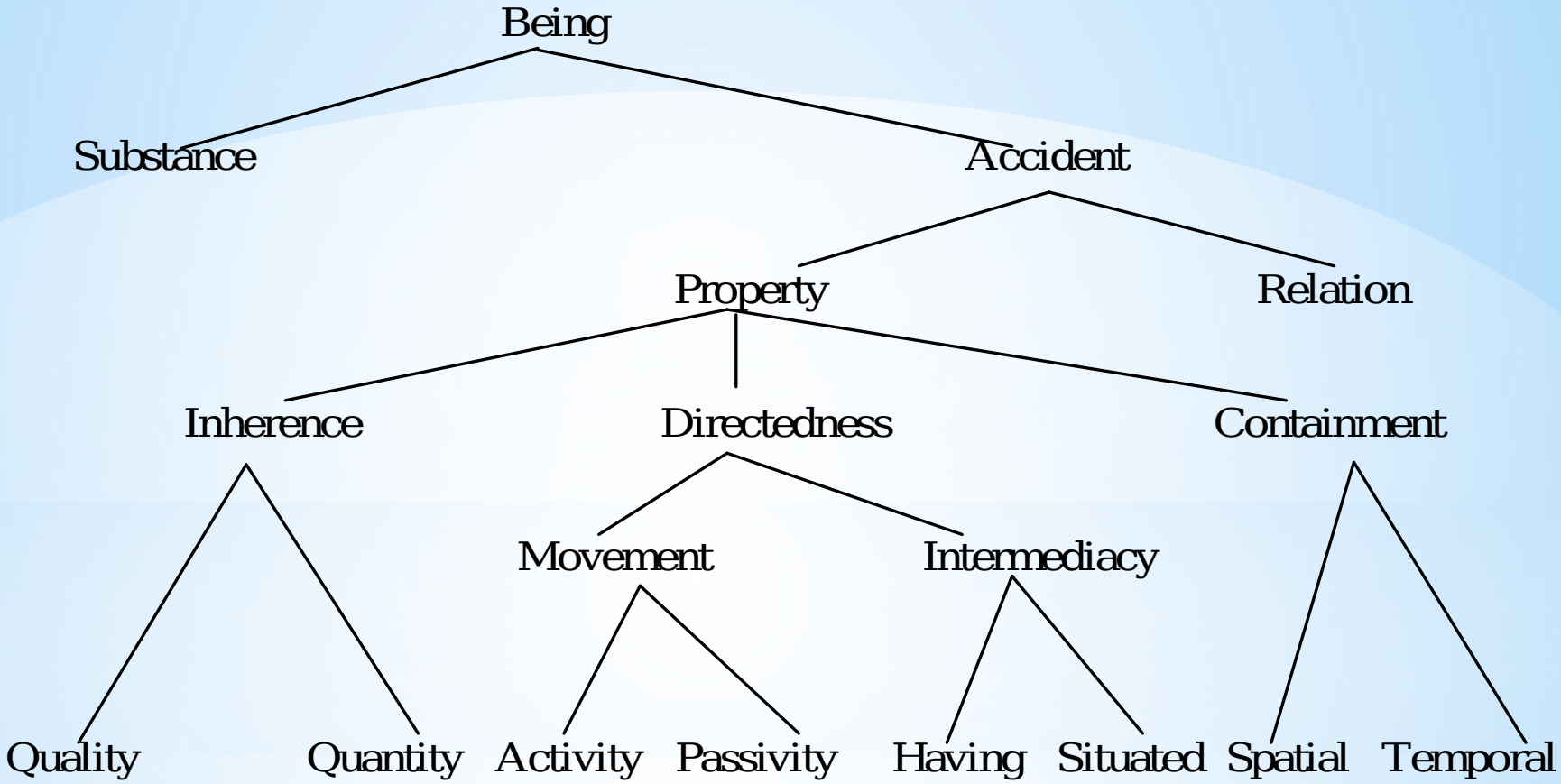
*Aristotle's work on natural history and logic laid out taxonomic principles.

* Taxonomy or the division of things into genera and species is a way of classifying predicates in the logic; it is a refinement of the ten basic categories of predicates.

* The basic idea was developed in detail in the long tradition of Aristotelian logic.

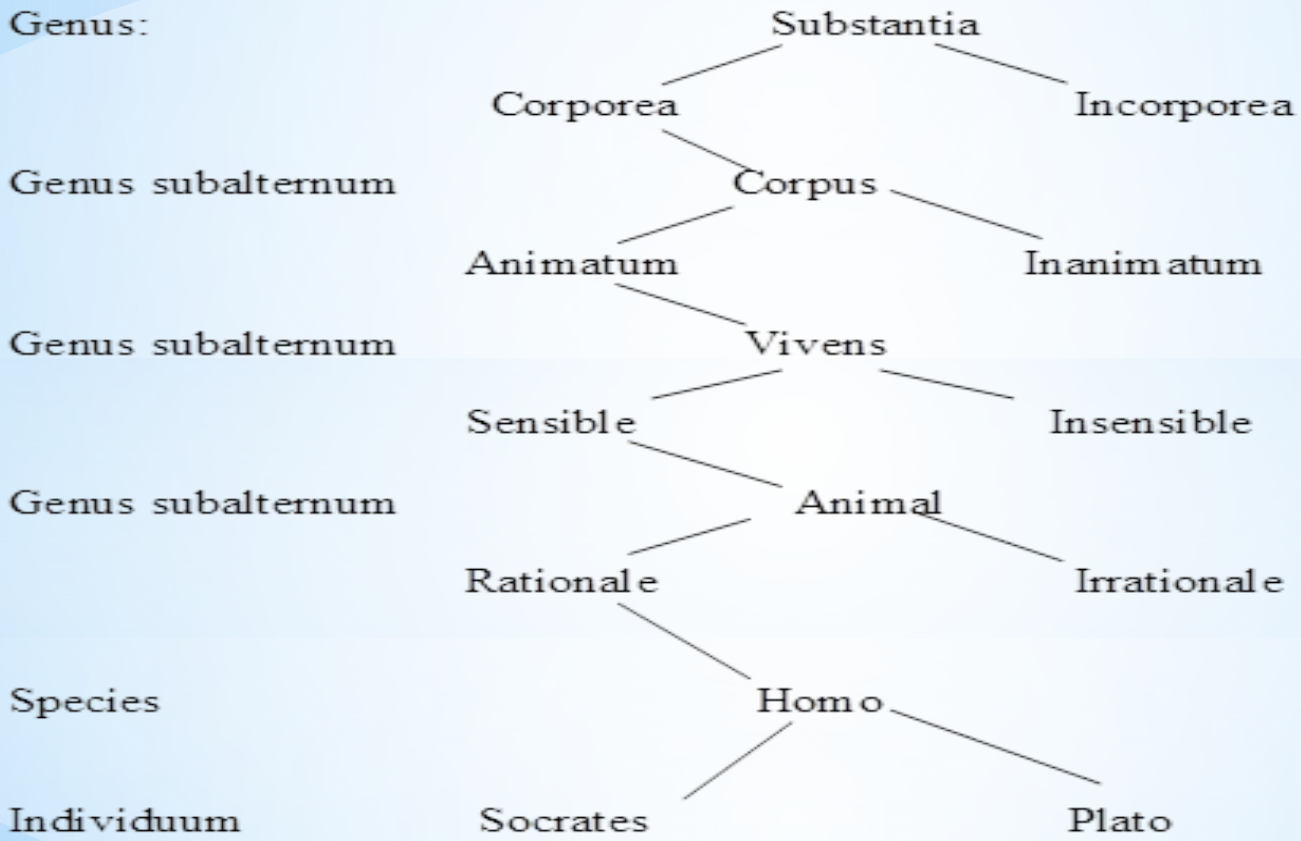
*Aristotle offered an ontology which included 10 categories, shown as the leaves in the tree shown in figure the given below (from Sowa, 2009, after Brentano).

* **Ontology of Aristotelian Origin**



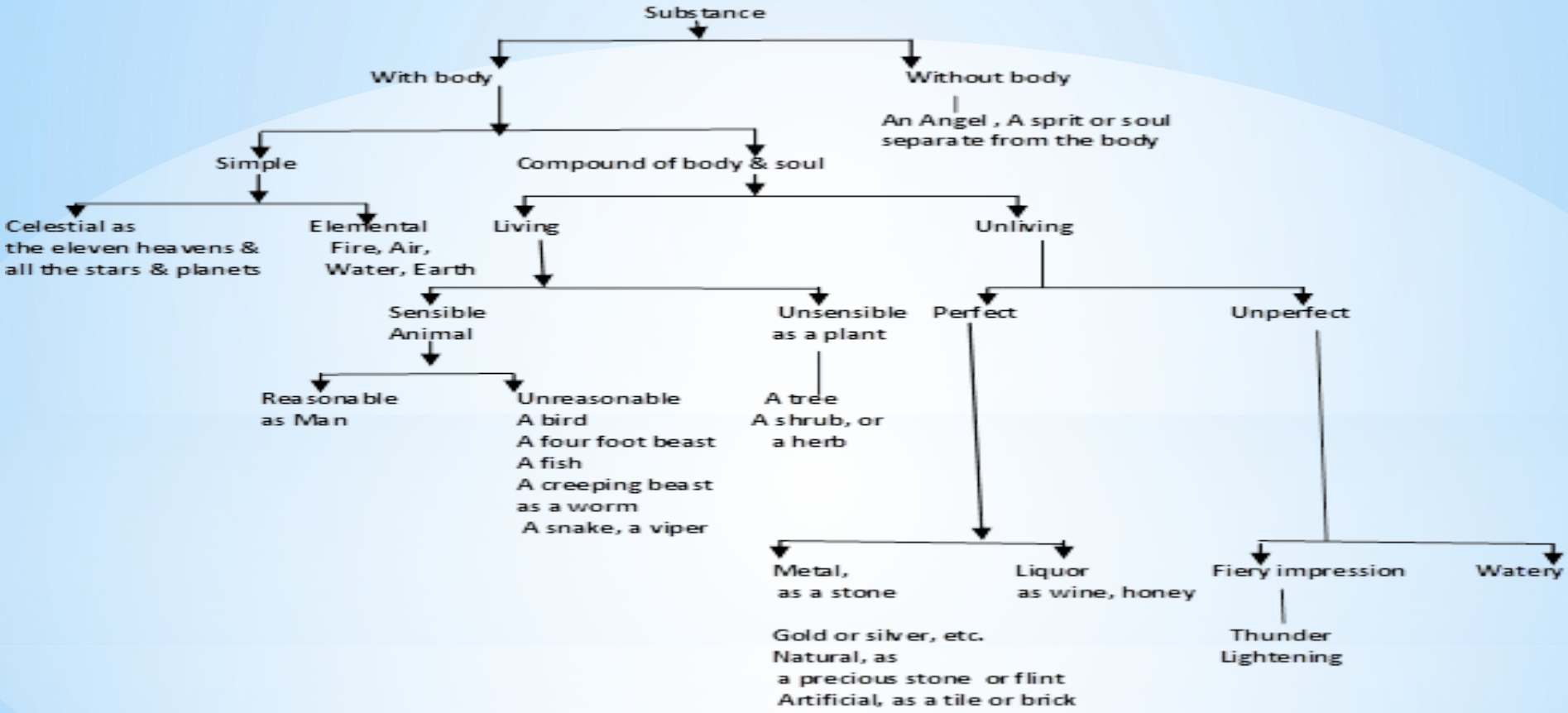
* **Ontology of Aristotelian Origin**

* Aristotle did not extend this division to develop a hierarchical arrangement, but his commentator Porphyry did and the Porphyrian tree of hierarchically linked genera and species became canonical in the tradition of Aristotelian logic. The tree provides a logical classification of the category of substance as the figure shows:



* Ontology of Aristotelian Origin

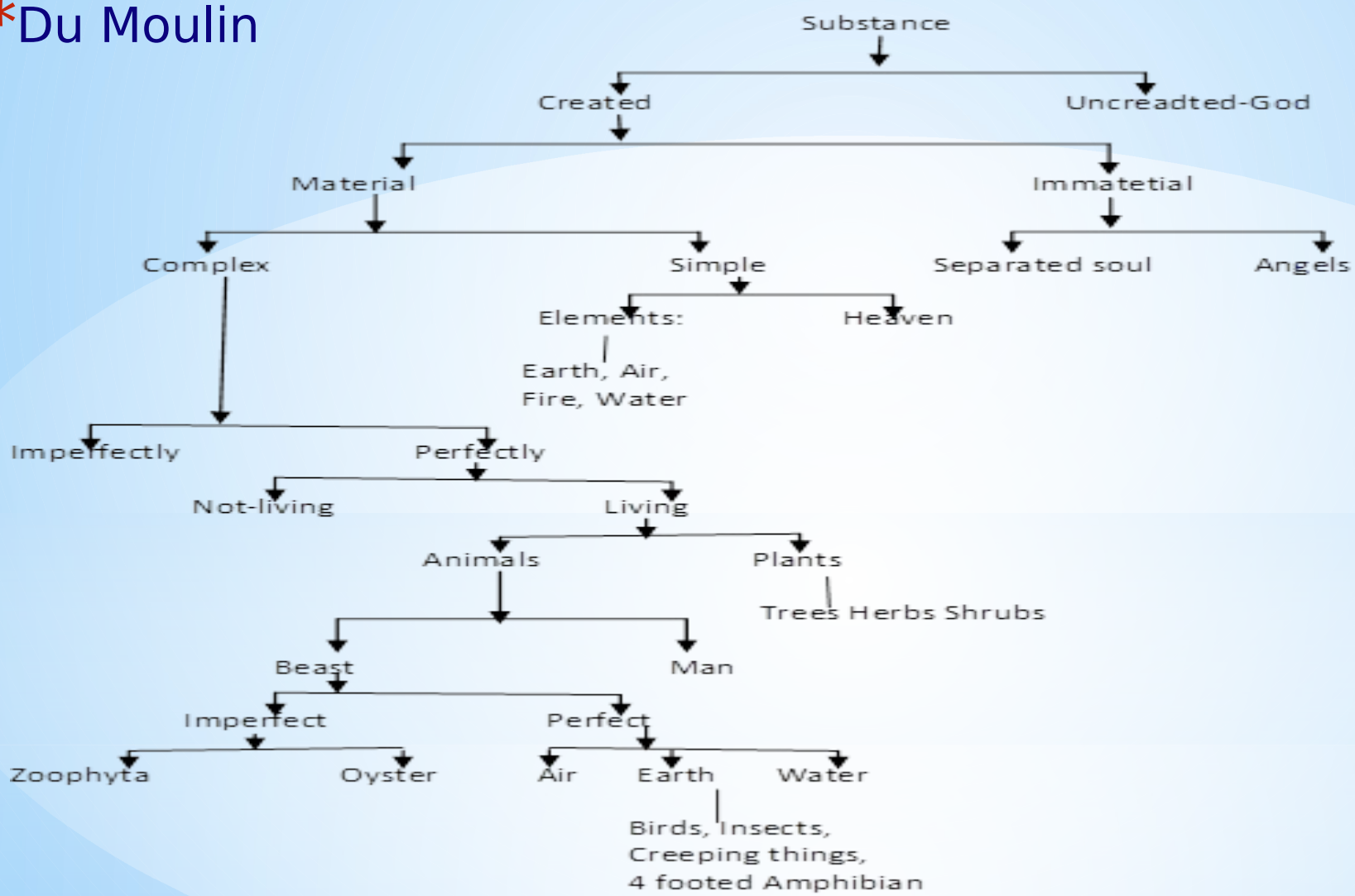
*Bunddeville



* **Ontology of Aristotelian Origin**

We find in Aristotelian logic books further divisions in each category and the philosophical and scientific lore of the age is included in these divisions and tables by the sixteenth and seventeenth centuries. The examples of Bunddeville and Du Moulin are typical of the encyclopedic nature of these logics (Slaughter, 1982:39).

*Du Moulin



* **Ontology of Aristotelian Origin**

*The development of a scientific (botanical) taxonomy.

* The following levels of taxa are found:

1. Unique Beginner: e.g. plant, animal

2. Life form: e.g. tree, bush, flower, weed, fern

3. Intermediate: this is an unstable category that manifests itself during a period of adjustment in the taxonomic system and then disappears when a settled, adjusted system is re-established

4. Genus: pine, oak, masterwort

5. Species: ponderosa pine, black oak

6. Variety: northern ponderosa pine, swamp white oak

* **Ontology of Aristotelian Origin**

* A Comparison of Berlin's folk taxonomy with a fully developed hierarchy of specialized taxa

FOLK TAXONOMY	SCIENTIFIC TAXONOMY
Unique Beginner	Kingdom
Life Form	Phylum
	Class
	Order
	Family
(Intermediate)	
Genus	Genus
Species	Species
	Sub-species
Variety	Variety

Specialized taxa

* **Ontology of Aristotelian Origin**

* A Comparison of Berlin's folk taxonomy with a fully developed hierarchy of specialized taxa

FOLK TAXONOMY	SCIENTIFIC TAXONOMY
Unique Beginner	Kingdom
Life Form	Phylum
	Class
	Order
	Family
(Intermediate)	
Genus	Genus
Species	Species
	Sub-species
Variety	Variety

Specialized taxa

* **Ontology in Thesaurus**

Plan of Classification in Roget's' Thesaurus

Class	Section	Sub section
ABSTRACT RELATIONS		I. Existence
		II. Relation
		III. Quantity
		IV. Order
		V. Number
		VI. Time
		VII. Change
		VIII. Causation
SPACE	I. In General	
	II. Dimension	
	III. Form	
	IV. Motion	
MATTER	I. In General	
	II. Inorganic	1. Solids
		2. Fluids
	III. Organic	1. Vitality
		2. Sensation
INTELLECT	I. Formation of Ideas	
	II. Communication of ideas	
VOLITION	I. Individual	
	II. Intersocial	1. In General
		2. Possessive relations
AFFECTIONS	I. In General	
	II. Personal	
	III. Sympathetic	
	IV. Moral	
	V. Religious	

Ontology in Thesaurus

Amarakosha

* The *Amarakosha* is a thesaurus of Sanskrit written by the ancient Indian scholar Amarasimha. It is the oldest extant *kosha*.
* The *Amarakosha* consists of verses that can be easily memorized.
* It is divided into three *khandas* or chapters. The first, *svargadi-khanda* ("heaven and others") has words pertaining to gods and heavens. The second, *bhuvargadi-khanda* ("earth and others") deals with words about earth, towns, animals and humans. The third, *samanyadi-khanda* ("common") has words related to grammar and other miscellaneous words.

Ontology in *naamatiipa nikaNTu* (Tamil)

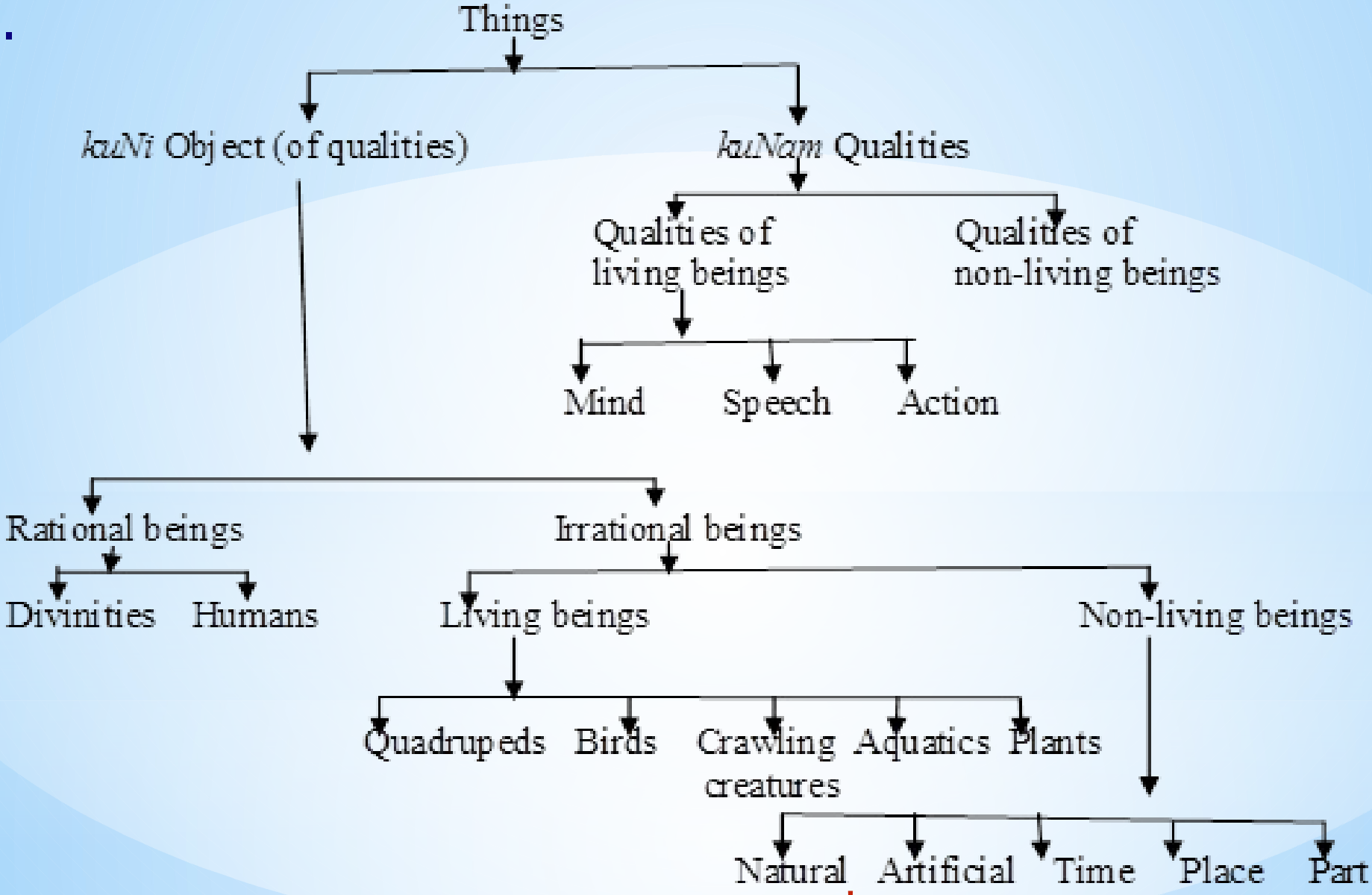
Rational creatures: 1. Divinities
2. Human Beings

Irrational creatures: 1. Quadrupeds
2. Birds
3. Crawling beings
4. Aquatics
5. Plants

Irrational non-living beings: 1. Natural things
2. Artificial things
3. Place
4. Time
5. Part

Qualities: 1. Living creatures: mental qualities
2. Living creatures: communicative qualities
3. Living creatures: qualities of action
4. Qualities of non-living beings

* **Ontology in *nikhantu***



* **Ontology in *nikhantu***

* Nida (1975) who was concerned with the preparation of a thesauric dictionary for Greek gives the following as the tentative hierarchical classification of the referential meanings or lexical concepts (Nida1975:178-186).
* His is a componential approach to meaning. He elaborately discusses about the foundation of his theory classification in his work entitled "Componential Analysis of Meaning: An Introduction to Semantic Structure" (Nida 1975) He has classified the lexical concepts under four categories: entities, events, abstracts and relationals.
* The following is the outline of his design (Nida, 1975: 178-186).

* Ontology in Nida's thesaurus dictionary

I. Entities

A. Inanimate

1. Natural

a. Geographical

b. Natural substances

c. Flora and plant products

2. Manufactured or constructed entities

a. Artifacts (non-constructions)

b. Processed substances: foods, medicines, and perfumes

c. Constructions

B. Animate entities

1. Animals, birds, insects

2. Human beings

3. Supernatural power or beings

* Ontology in Nida's thesaurus dictionary

II. Events

A. Physical, B. Physiological, C. Sensory, D. Emotive, E. Intellection, G. Communication, G. Association, H. Control, I. Movement, J. Impact, K. Transfer, L. Complex activities, involving a series of movements or actions

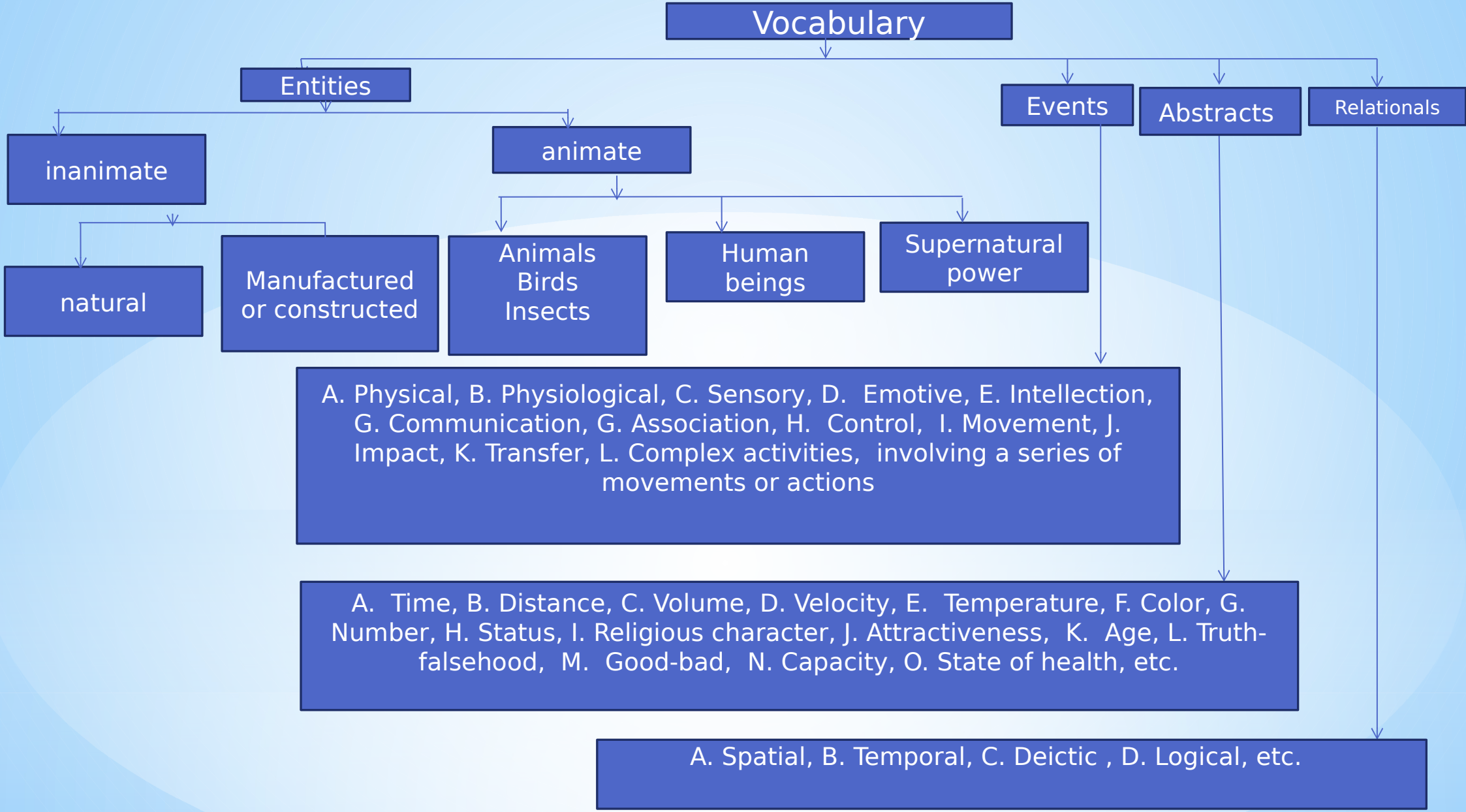
III. Abstracts

A. Time, B. Distance, C. Volume, D. Velocity, E. Temperature, F. Color, G. Number, H. Status, I. Religious character, J. Attractiveness, K. Age, L. Truth-falsehood, M. Good-bad, N. Capacity, O. State of health, etc.

IV. Relationals

A. Spatial, B. Temporal, C. Deictic , D. Logical, etc.

* **Ontology in Nida's thesaurus dictionary**


Vocabulary
Entities
Events
Abstracts
Relationals
inanimate
animate
natural
Manufactured or constructed
Animals Birds Insects
Human beings
Supernatural power
A. Physical, B. Physiological, C. Sensory, D. Emotive, E. Intellection, G. Communication, G. Association, H. Control, I. Movement, J. Impact, K. Transfer, L. Complex activities, involving a series of movements or actions
A. Time, B. Distance, C. Volume, D. Velocity, E. Temperature, F. Color, G. Number, H. Status, I. Religious character, J. Attractiveness, K. Age, L. Truth-falsehood, M. Good-bad, N. Capacity, O. State of health, etc.
A. Spatial, B. Temporal, C. Deictic , D. Logical, etc.

*** Ontology in Nida's thesaurus dictionary**

* Thesaurus is a in its wider sense is a classification of words by concepts, topics, or subjects.
* The present Tamil Onto-thesaurus is the extended version of Electronic thesaurus of Tamil focusing more on the ontological features.
* Two kinds of issues arise in the preparation of Tamil onto-thesaurus:
  * Linguistic issues
  * Computational issues

* **Tamil Visual Onto-thesaurus**

Linguistic Issues

It involves mainly the following four tasks:

1. Developing ontology for Tamil based on structural semantic principles.

2. Establishing semantic domains and sub domains based on distinguishing semantic or componential features of lexical items.

3. Classifying Tamil vocabulary to fit into the ontology developed.

4. Linking words by various semantic or lexical relations such as synonymy, hyponymy-hyperonymy, meronymy-holonymy, compatibility, and incompatibiliity.

* **Tamil Visual Onto-thesaurus**

Computational Issues

It involves mainly the following three tasks:

* Conversion of linguistic data base into computer accessible format.
* Preparation of a tool to provide the facilities for augmenting, entering and editing the raw data, and classifying the lexical items in a semi-automatic way.
* Creation of user friendly interfaces for accessing the onto-thesaurus in simple manner.

* **Tamil Visual Onto-thesaurus**

* Ontology of Tamil Vocabulary

* The ontology available in Rajendran (1982, 2001), which is founded on the theory of componential analysis of meaning propounded by Nida (1975a), Indian tradition of *nikhandu* and Aristotelian principle of genera and species is enhanced to suit the present purpose.
* Following Nida (1975a) the vocabulary of Tamil is grouped initially into four domains: entities which consist of referential meanings of concrete concepts, events which consist mainly of verbs and verbal nouns and abstracts which consist mainly of adjectives and adverbs apart from abstract nouns and relationals which consists of functional 'words including postpositions, connectives and coordinators.
* These four domains are further hierarchically classified into sub-domains under which the lexical items are listed.
* One will be able to capture the meaning of the concerned lexical item from the domain to which it belongs in a hierarchical fashion.
* A full-fledged detail of building ontology for the vocabulary of a language is available in Nida (1975a).

Structuring of Vocabulary by Lexical Relations

* Lexical semantics offers foundation for structuring vocabulary in terms of lexical relations (Lyons 230-335, Cruses 1986).
* In the NLP oriented papers, the general practice are to avoid giving linguistic details based on which the system is built.
* But here we would like to give the lexical semantics of building onto-thesaurus to make it more transparent.

* **Ontology of Tamil Vocabulary**

Congruence Relations

*There are four basic relations between classes that furnish a model for establishing the fundamental group of sense relations and for defining a set of systematic variants applicable to virtually all other paradigmatic sense relations; they are identity, inclusion, overlapping, and disjunction (Cruse, 1986:86-87).

Identity: class A and class B have same members.

Inclusion: class B is wholly included in the class A

Overlap: class A and class B have members in common but each has members not found in the other

Disjunction: class A and class B have no members in common

*These four congruence relations culminate into the four lexical relations discussed below.

Lexical Relations

* There are at least four lexical or meaning relations by which lexical items can be linked or related to one another in the ontological structure of Tamil vocabulary.
* They are synonymy, hyponymy, compatibility and incompatibility (Lyons, 1977; Cruse, 1986:84-111).
* A word acquires its referential meaning in being a member of a semantic domain by the common features it shares with other members in that domain, and having contrasting features which separate it from other members of the domain.
* It is the semantic relations among words, such as synonymy, hyponymy, compatibility and incompatibility, which help one to classify and organize words in terms of semantic domains in a structural fashion.

* **Ontology of Tamil Vocabulary**

* Synonymy: e.g. *puttakam* 'book': *nuul* 'book';
* Hyponymy-hypernymy: e.g. *pacu* ‘cow’ : *vilangku* 'animal’;
* Meronymy-holonymy: e.g. *uTal* ‘body’: *kaal* ‘leg’;
* Compatibility: e.g. *cellappiraaNiv* 'pet': *naay* 'dog'
* Incompatibility: *malai* ‘hill’: *maTu* ‘water hole’; incompatibility leads to the relation called opposition which culminates into many types which are discussed below.

*** Ontology of Tamil Vocabulary**

Lexical Inheritance

Hypernymy-hyponymy and meronymy-holonymy assure a lexical item to inherit semantic features as exemplified below:

kuiyl 'quail' – *iniya kuraluTaiya kariya paRavai* 'a black bird with sweet voice'

paRavai – irukaalkaLum alakukaLum uTaiya, uTalin iruppakkangkaLilum paRappataRku eeRRavakaiyil ciRakukaL koNTa vilangkinam 'an animal with two legs and a beak and wings at its sides of its body for flying'

vilangkinam 'animal' – *taanaaka iyangkum uTal uRuppukaLum celluloos ilaata celkaLum uLLa uyiringkaL'* living beings with parts functioning automatically and cell walls without cellulose'

uyirinam - uyir vaazkiRa onRu 'one which lives'

Lexical Oppositions

* There are many types of oppositions (Lyons, 1977: 270-290; Cruse, 1986: 197-263, Cruse, 165-176).
* They can be grouped into two types of oppositions or contrasts based on the number of items involved in the contrast: binary contrast or opposition and non-binary contrast or non-binary opposition.
* If the contrast is made between two lexical items, it is called binary contrast; if the contrast is made between more than two lexical items, it is called non-binary contrast. The binary-contrasts are listed below with examples from Tamil.

* **Gradable opposites**: The gradable antonyms or opposites occur as end points on a scale, and the denial of one member of the pair does not imply the assertion of the other. E.g. *nalla* 'good': *keTTa* 'bad'; *azakaana* ‘beautiful’ : *kuruuramaana* ‘ugly’, *uyaramaana* ‘long’ : *kuLLamaana* ‘short’.
* **Ungradable opposites or non-gradable opposites**: Non-gradable antonymy divides the universe of discourse absolutely without admitting degrees of more or less. E.g. *aaN* 'male': *peN* 'female'; *kalyaaNamaana* ‘married’: *kalyaaNamaakaata* ‘un-married’, *uyiruLLa* ‘live’: *uyiraRRa* ‘dead’.
* **Complementaries:** Complementaries do not fall on a scale, though in certain borderline cases something like a scale might exist. E.g. *pakal* 'day: *iravu* 'night', *uNmai* 'truth': *poy* 'false', *aaN* 'male': *peN* 'female' *veRRi* 'success': *toolvi* 'failure'.
* **Privative Opposites**: In privative opposition, one item of the pair denotes some positive property and the other denotes the absence of that property. They are referred as privative opposites. E.g. *uyiruLLavai* 'living beings': *uyirillaatavai* 'non-living beings'.
* **Equipollent opposites**: In equipollent opposition, both the contrasting lexemes denote a positive property. They are referred as equipollent opposites. E.g. *aaN* 'male': *peN* 'female'.

* Ontology of Tamil Vocabulary

Converses: Converseness is the relationship that holds between such pairs of words (i.e. converses) in which one can be considered as the reverse of the other. Converseness is distinguished from antonymy and complementarity. e.g, *kaNavan* 'husband': *manaivi* 'wife', *kol* 'kill': *kollappaTu* 'be killed', *vaangku* 'get': *koTu* 'give'; 'X is Y's husband' means 'Y is X's wife'. There are various types of converses.

1. Converse pairs of social roles, e.g. *vaitiyar* 'doctor' : *nooyaaLi* 'patient', *ejamaanar*/*ejamaani* 'boss': *veelaiyaaL* 'servant';

2. Converse pairs of Kinship terms, e.g. *tandtai* 'father'/*taayaar* 'mother' : *makan* 'son'/*makaL* 'daughter';

3. Converse pairs of temporal relations, e.g. *munnaal* before': *pinnaal* 'after', *munnar* earlier': *pinnar* 'later';

4. Converse pairs of spatial relations, e.g. *munnaal* 'in front': *pinnaal* 'at the back', *meelee* 'above': *kiizee* 'below'

* **Ontology of Tamil Vocabulary**

* **Directional opposites**: e.g. *meelee* 'above' : *kiizee* 'below', *vandtuceer* 'arrive': *puRappaTu* 'start', *vaa* 'come': *poo* 'go'
* **Orthogonal opposites** (perpendicularly opposites**)**: e.g *vaTakku* 'north' : *meeRku* 'west', *vaTakku*: *kizakku* 'east'; i.e. *meeRku* 'west' and *kizakku* 'east' are perpendicularly opposite to *vaTakku* 'north'
* **Antipodal opposites** (diametrically opposites): *vaTakku* 'north' : *teRku* 'south'; *kizakku* 'east' : *meeRku* 'west'; i.e. *vaTakku* 'north' is diametrically opposite to *teRku* 'south' and *kizakku* 'east' is diametrically opposite to *meeRku* 'west'

* **Ontology of Tamil Vocabulary**

Non-binary Opposition:

*As stated above, if the opposition involves more than two lexical items such contrasts are called non-binary opposition (Lyon, 1977:271).

*There a number of types of non-binary contrasts.

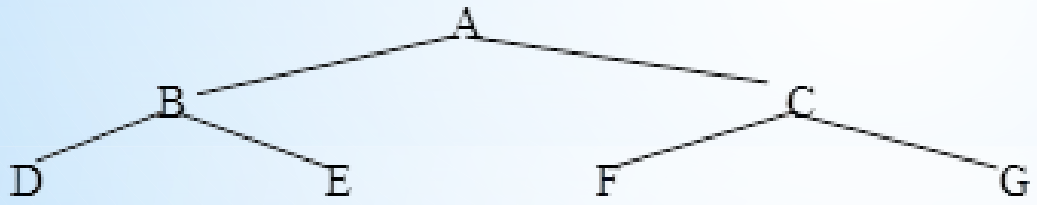
*They are dealt under non-branching hierarchies.

Hierarchies

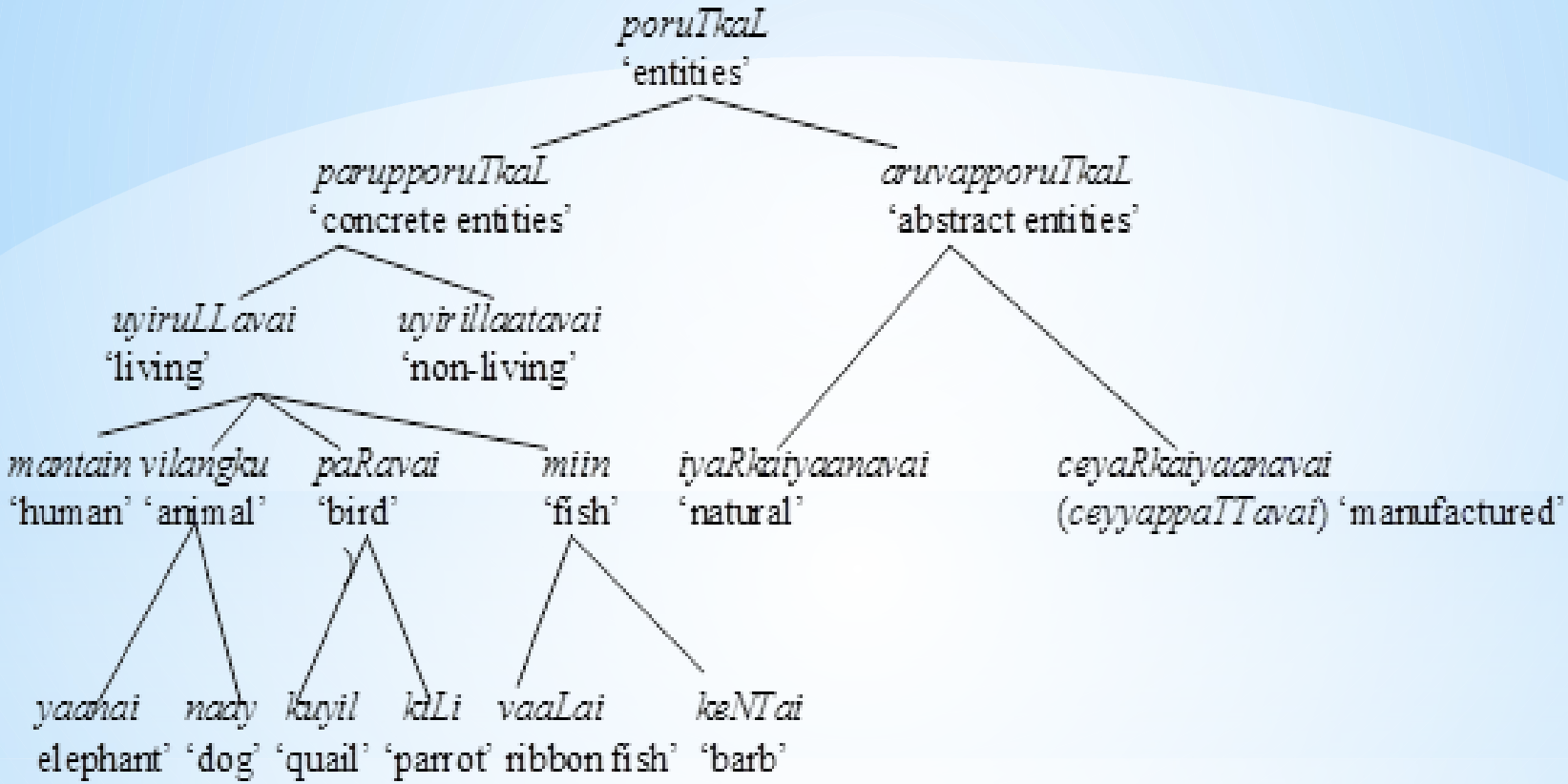
Hierarchies (Cruse 1986:112-118) are of two types: branching hierarchies and non-branching hierarchies.

The branching hierarchy shows tree structure, where as non-branching hierarchy does not show tree structure.

. Branching hierarchies

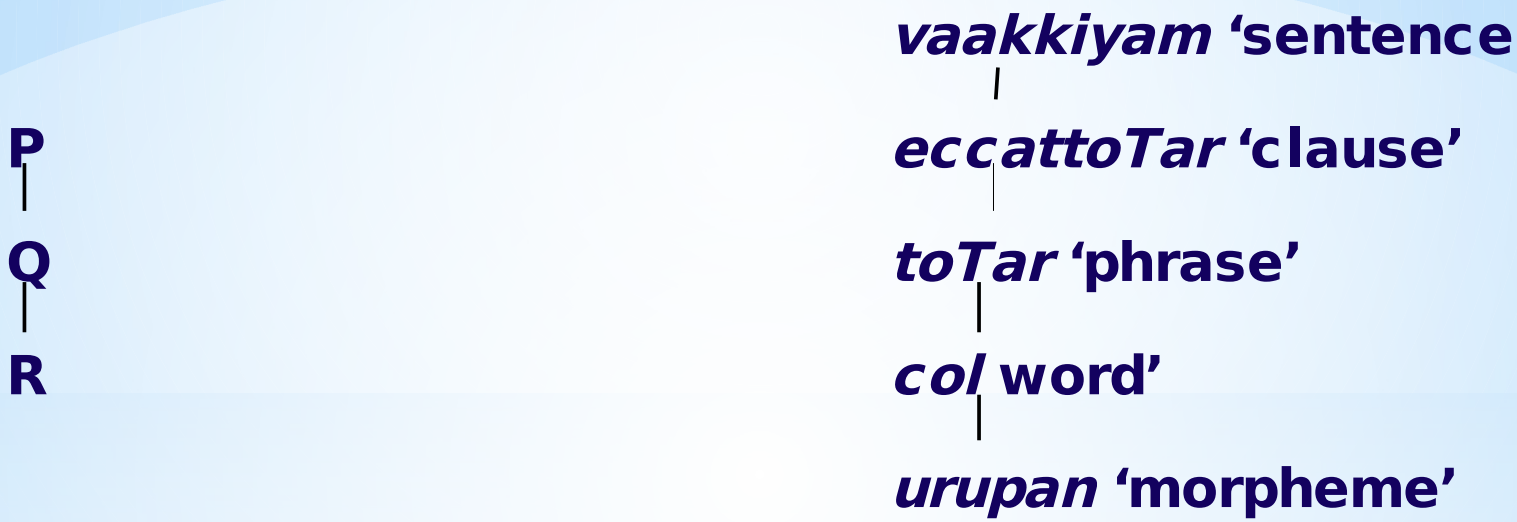


Branching hierarchies



*** Ontology of Tamil Vocabulary**

Non-branching hierarchies



* **Ontology of Tamil Vocabulary**

Taxonomic Hierarchies

* The first major type of branching lexical hierarchy is the outcome of the hyponymy-hypernymy relation between lexical items.

* The sequence of hyponymy-hypernymy relation leads to branched-hierarchical structure of a set of vocabulary items which show this pair of relations among themselves.

* Taxonomic hierarchies (Cruse, 1986: 136-155) are more liberal than hyponymy-hypernymy hierarchies. The following is an example.

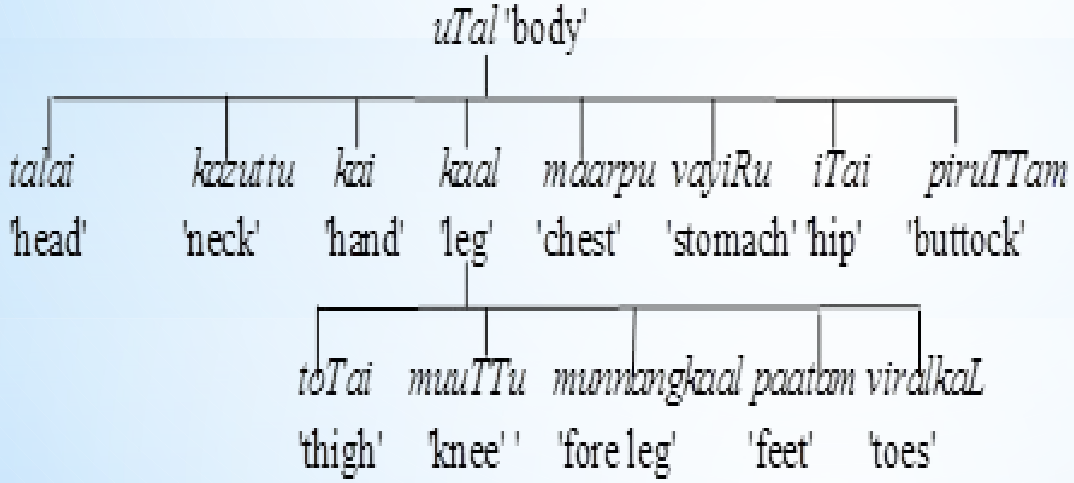
* **Ontology of Tamil Vocabulary**

Taxonomic Hierarchies

* **Ontology of Tamil Vocabulary**

Meronymic Hierarchies

* The second major type of branching lexical hierarchy is the part-whole type which is called meronomies. (Cruse, 1986: 157-180).
* Meronymic hierarchies are the result of meronymy-holonymy relation shown by the lexical items.
* The meronymy-holonymy relation also gives hierarchical structure to a set of vocabulary. The following is an example.



* **Ontology of Tamil Vocabulary**

Non-branching Hierarchies

Non-binary opposition leads to a number of types of non-branching hierarchies. They are listed below:

Bipoles: The bipoles is the simplest kind of linear structure found in a pair of opposites. They are simply oppositions which we have discussed earlier (Cruse, 2000: 189). e.g. *niiLamaana* ‘long’: *kuTTaiyaana* ‘short’, *viraivaaka* ‘fast’: *metuvaaka* ‘slowly’.

Bipolar chains: The bipolar chains imply a scale on which a pair of opposites operates denoting different degrees of the property. They have implicit superlative terms of opposite polarity at each end of the scale (Cruse, 2000: 189). The following is the example: *mikanuNNiya* 'very minute', *nuNNiya* 'minute', *ciRiya* 'small', *periya* 'big' *mikapperiya* 'very big'

Monopolar Chains: In monopolar, there is no sense that the terms at the ends of the chain are oriented in opposite directions. Degree, stages, measures, ranks and sequences come under monopolar chains (Cruse , 2000: 190).

Degrees: Degrees incorporate as part of their meaning different degrees of some continuously scaled property such as size or intensity (Cruse, 2000: 190). e.g**.** *paaRai* 'mound': *kunRu* 'hillock': *malai* 'hill' : *maamalai* 'mountain'

Stages: Stages are points in a life cycle of something and normally involve the notion of progression (Cruse 2000: 190). e.g. *kuzandtaip pruvam* 'child stage': *vaalipap paruvam* 'young stage' : *mutir paruvam* 'adult stage': *mutu paruvam* 'old stage'.

Measures: Measures are based on part-whole relationship, with each whole divided into a number of identical parts (Cruse, 2000: 190). *cekaNTu* 'second', *miniT* 'minute', *maNi* 'hour', *naaL* 'day', *vaaram* 'week' *maatam* 'month', *aaNTu* 'year'.

Ranks: The lexical items under ranks entails a sequential order which is not gradual (Cruse, 2000. 191). e.g. *virivuraiyaaLar* 'lecture', *mutunilai virivuraiyaaLar* 'senior lecture', *iNaippeeraaciriyar* 'reader', *peeraaciriyar* 'professor'. Lyons (1977: 290) includes numeral under rank as an unique type. E.g. *onRu* 'one', *iraNTu* 'two' .. *pattu* 'ten', *irupatu* 'twenty, ... *nuuRu*, *iRunuuRu*, ...

Sequences: Sequences are lexical items which are ordered but do not bear increasing property as in the case of previous chains (Cruse, 2000: 191). e.g. *iRandta kaalam* 'past tense', *nikaz kaalam* 'present tense', *etir kaalam* 'future tense'; *kaar kaalam* 'rainy season' *kuuLir kaalam* 'cold season', *munpani kaalam* 'early cold season', *pinpanik kaalam* 'late cold season', *iLaveeniR kaalam* 'milder hot season', *mutuveeniR kaalam* 'hot season'.

* **Ontology of Tamil Vocabulary**

Cyclical Sets or Cycles:

*The sequential lexical item (Cruse, 1986:187-190) can entail a cyclical order of time in the natural arena.

*e.g. *kaarkaalam* 'rainy season' *kuutirkaalam* 'cold season', *munpanikkaalam* 'early cold season', *pinpanikkaalam* 'late cold season', *iLaveenirkaalam* 'milder hot season', *mutuveenirkaalam* 'hot season';

* *canavari* 'January', *pepravari 'February'*, *maarc* 'March', *eeppiral* 'April', *mee* 'May', *juun* 'June'...};

njaayiRu 'Sunday', *tingkaL* 'Monday', *cevvaay* 'Tuesday' , *putan* 'Wednesday', *viyaazan* 'Thursday', *veLLi* 'Friday', *cani* 'Saturday'.

Propositional series or grids

pacu 'cow': *kanRu* 'calf' :: *kutirai* 'horse': *kuTTi* 'foe' ;

paampu 'snake' : *kunjcu* 'young one':: *palli* 'lizard' : *kunjcu* 'young one' ;

yaanai : *piLiRu 'trumpet'* :: nari ‘fox’: *uuLaiyiTu* 'howl'.

Propositional series or grids help us to relate lexical items by different types of meaning relations.

Even derivative relation can be expressed using grids;

e.g. *kuLi* ‘bathe’ : *kuLittal* ‘bathing:: *kuTi* ‘drink’:: *kuTittal* ‘drinking’.

* As mentioned earlier, the lexical items are arranged into four major categories entities, events, abstracts and relationals in line with Nida (1975a).
* Each category requires different representation in the VOT because of their inherent componential features.
* The organization of lexical items in the data base is discussed below:

* **Creation of database for VOT**

* Generating ontology tree from a given huge database for a user's query is the backbone activity for the Tamil onto-thesaurus which an intelligent information retrieval and visualization system. This type of tool provides a powerful way for data representation and knowledge mining process.
* The user can input search word to the system and extract the relevant information form the database. The information is given in a visual representation.
* A well-structured data base is need for this purpose. The following table gives the different types of meaning relations established between words or lexical items and the unique way of representing them in the database for the ontology tree generation.

* BUILDING OF ONTO-THESAURUS

Meaning Relations	Their Tamil Equivalents	Relation distinguishing Symbols used in the data base
Synonymy	□□□□□□□□□□	<>
Hyponymy	□□□□□□□□□□	{ }
Meronym	□□□□□□□□□□□□	/ |
The relation between the epicene noun and the gender marked nouns.	□□□□□□□□□□□	$#
Antonymy	□□□□□□□□□□□□	*
The relation between the verb and the nouns derived from the verb.	□□□□□□□□□□□□□□□	~ `
The relation between a noun and the modifiers (adjectives and adverbs) derived from the noun	□□□□□□□□□□	^ !

* BUILDING OF ONTO-THESAURUS

Relations	Subtypes	Example
Synonymy		*puttakam* 'book' to *nduul* 'book'
Hypernymy-Hyponymy		*vilangku* 'animal' to *paaluuTTi* 'mammal'
Hyponymy-Hypernymy		*pacu* 'cow' to *paaluuTTi* 'mammal'
Holonymy-Meronymy	Wholes to parts	*meecai* 'table' to *kaal* 'leg'
,,	Groups to members	*tuRai* 'department' to *peeraaciriyar* 'professor'
Meronymy-Holonymy	Parts to wholes	*cakkaram* 'wheel' to *vaNTi* 'cart'
,,	Members to groups	*paTaittlaivar* 'captain' to *paTai* 'army'

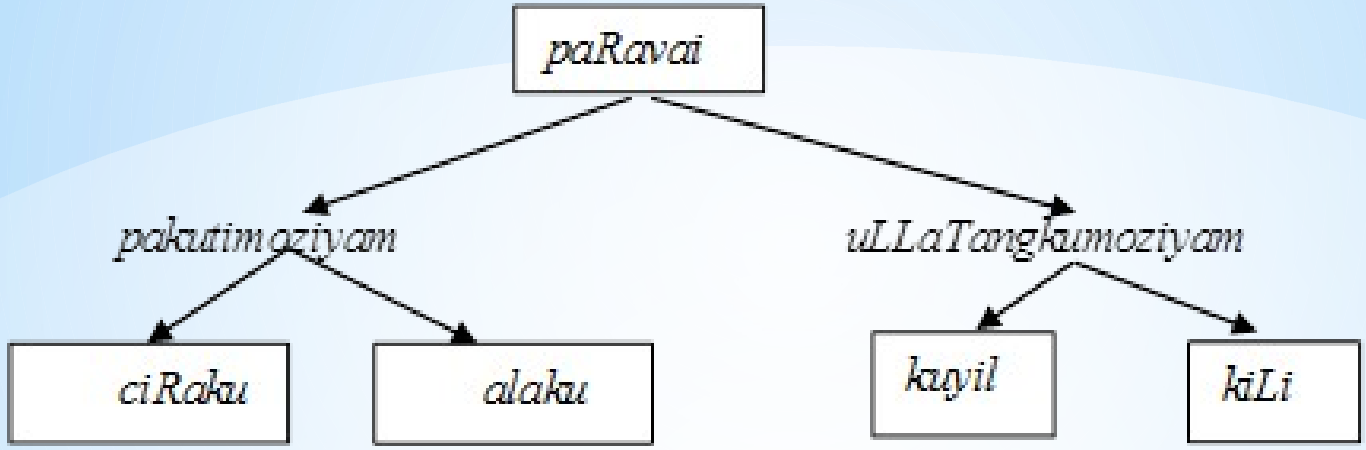
* Organization of entities in VOT

Binary Opposites	Antonymic (gradable)	*ndallavan* 'good person' to *keTTavan* 'bad person'
,,	**Complementary**	*aaTavar* 'man': *makaLir* 'woman'
	Converse	*kaNavan* 'husband' *to manaivi* 'wife'
,,	Privative (opposing features)	*ahRiNai* 'irrational' to *uyartiNai* 'rational'
,,	Equipollent (positive features)	*aaN* 'male' to *peN* 'female'
,,	Reciprocal Social roles	*vaittiyar* 'doctor' to *ndooyaaLi* 'patient'
,,	Kinship Relations	*ammaa* 'mother' to *makaL* 'daughter'
,,	**Antipodal Opposition**	*cikaram* 'peak' to *aTi* 'foot (of mountain)'
	Orthogonal opposition	*ciRumi* 'girl' : *ciRuvan* 'boy' *and peNTir* 'woman'

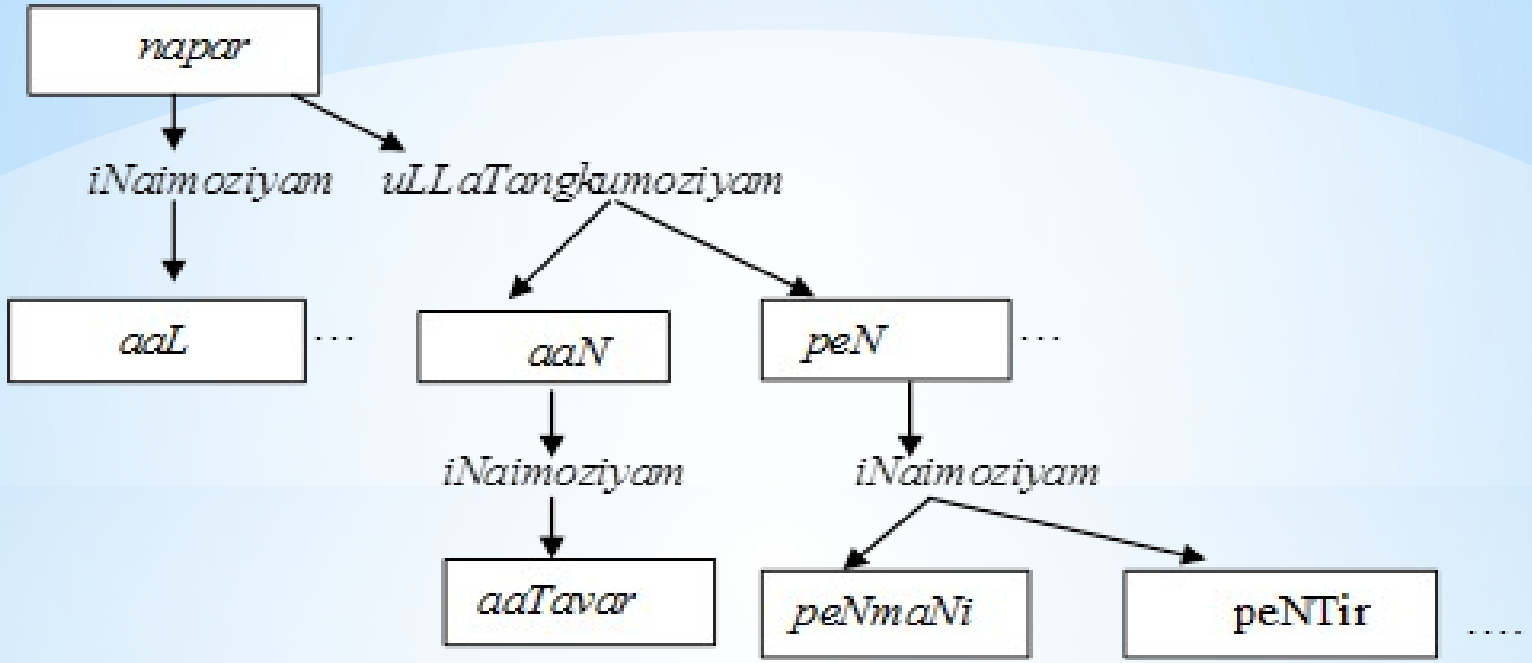
* Organization of entities in VOT

“	**Degrees**	*paaRai* 'mound': *kunRu* 'hillock': *malai* 'hill' : *maamalai* 'mountain'
,,	**Ranks**	*virivuraiyaaLar* 'lecture', *mutunilai virivuraiyaaLar* 'senior lecture', *iNaippeeraaciriya* 'reader', *peeraaciriyar* 'professor'
Compatibility		*ndaay* ‘dog’ to *cellappiraaNi* ‘pet’

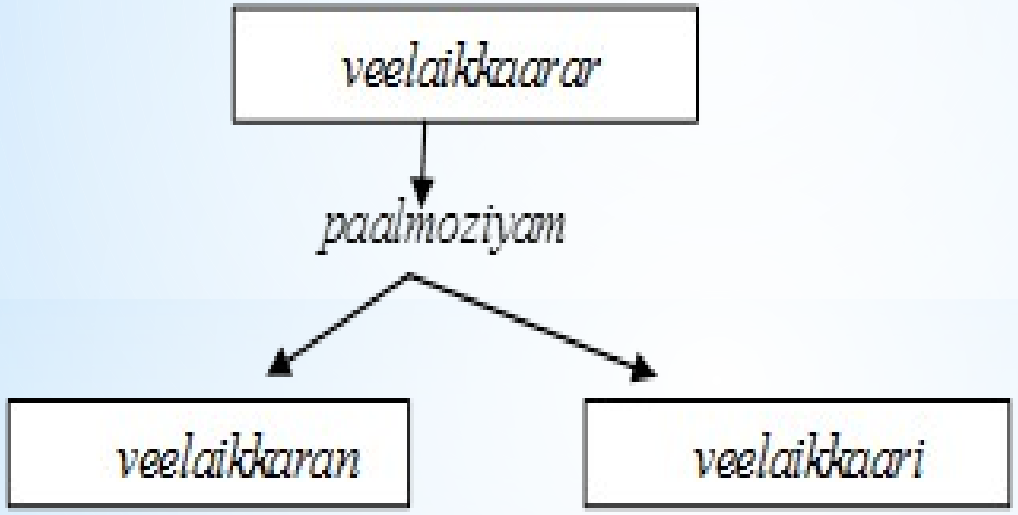
* Organization of entities in VOT



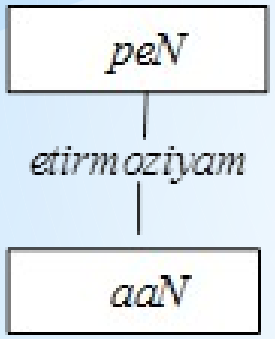
* Organization of entities in VOT



* Organization of entities in VOT



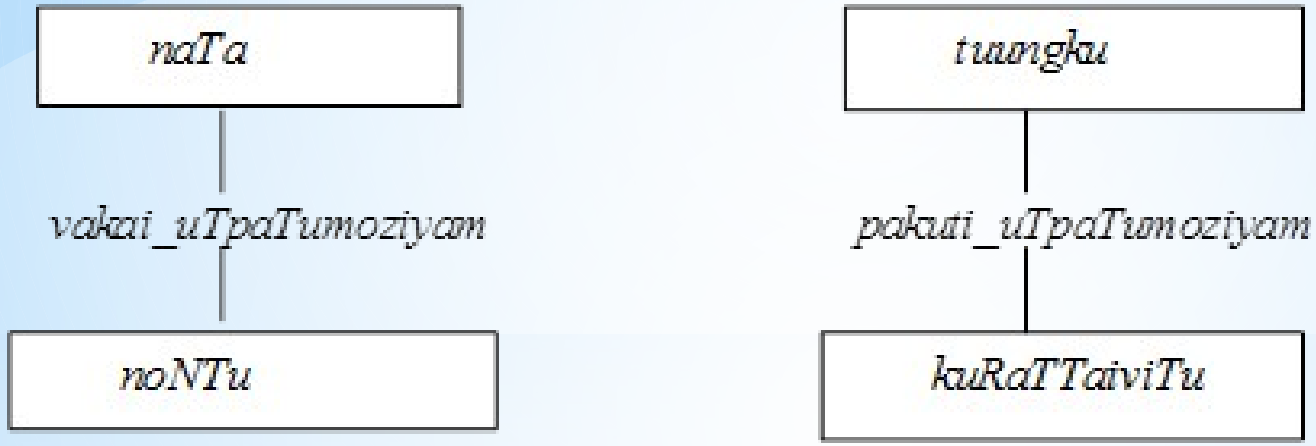
* Organization of entities in VOT



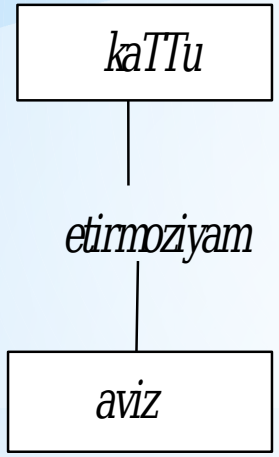
* Organization of entities in VOT

Relations	Definition/sub types	Example
Synonymy	Replaceable events	*tuungku* 'sleep' → *uRangku* 'sleep'
Meronymy-Hypernymy	From events to superordinate events	*paRa* 'fly' → *pirayaaNi* 'travel'
Troponymy	From events to their subtypes	*naTa* → *noNTu* 'limp'
Entailment	From events to the events they entail	*kuRaTTaiviTu* 'snore' → *tuungku* 'sleep'
"	From event to its cause	*uyar* 'rise' → *uyarttu* 'raise'
"	From event to its presupposed event	*vel* 'succeed' → *muyal* 'try'
"	From even to implied event	*kol* 'murder' → *iRa* 'die'
Antonym	Opposites	*kuuTu* 'increase' → *kuRai* 'decrease'; *kaTTu* 'build' *iTi* 'demolish'
"	Conversensess	*vil* 'sell' → *vaangku* 'buy'
"	Directional opposites	*puRappaTu* 'start' → *vandtuceer* 'reach'
Derivatives	Verb to verbal noun	*paTi* 'study' → *paTippu* 'education'

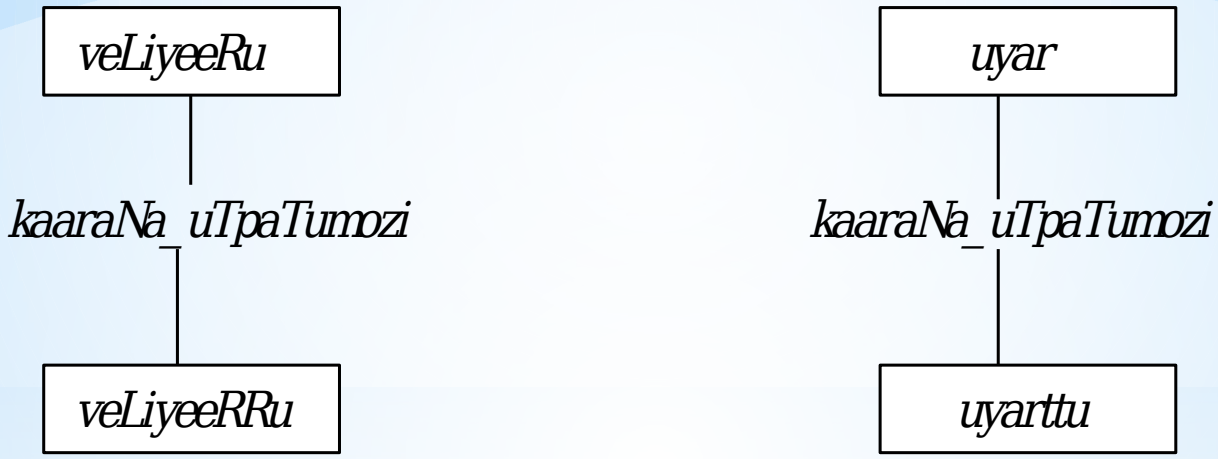
* Organization of Events in VOT



* Organization of events in VOT



* Organization of events in VOT



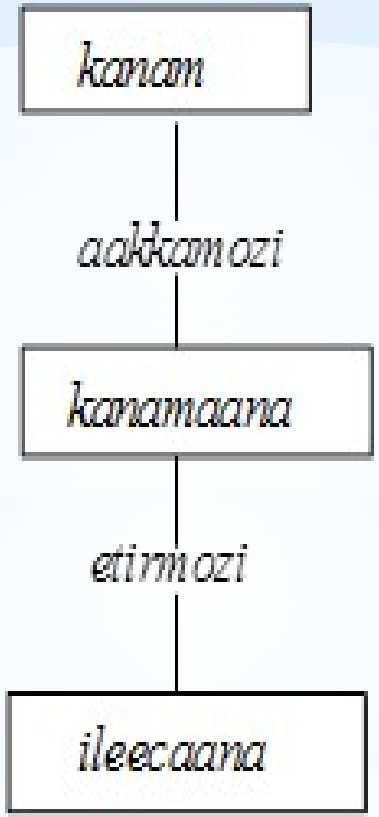
* Organization of events in VOT

Relations	Subtypes	Example
Synonymy		*tukkam* 'sorrow' to *tunpam* 'sorrow', *cangkaTam* 'sorrow', *tuyaram* 'sorrow'
Hypernymy-Hyponymy		*uNarcci* 'feeling' to *makizcci* 'happiness'
Hyponymy-Hypernymy		*paccai* 'green' to *niRam* 'colour '
Holonymy-Meronymy	Wholes to parts	*vaaram* 'week' to *naaL* 'day'
Meronymy-Holonymy	Parts to wholes	*injcu* 'inch' to *aTi* 'feed'
Binary Opposites	Antonymic (gradable)	*nalla* 'good' to *keTTa* 'bad'
,,	Temporal Relations	*munnar* 'before' to *pinnar* 'after'
,,	Orthogonal or perpendicular opposition	*vaTakku* 'north' to *kizakku* 'east' and *meeRku* 'west'
,,	**Antipodal Opposition**	*vaTakku* 'north' to *teRku* 'south'
Multiple opposites	**Serial**	*onRu* 'one', *iraNTu* 'two', *muunRu* 'three', *ndaanku* 'four'
,,	**Cycle**	*njaayiRu* 'Sunday' to *tingkaL* 'Monday' .. to *cani* 'Saturday'

* Organization of Abstracts in VOT

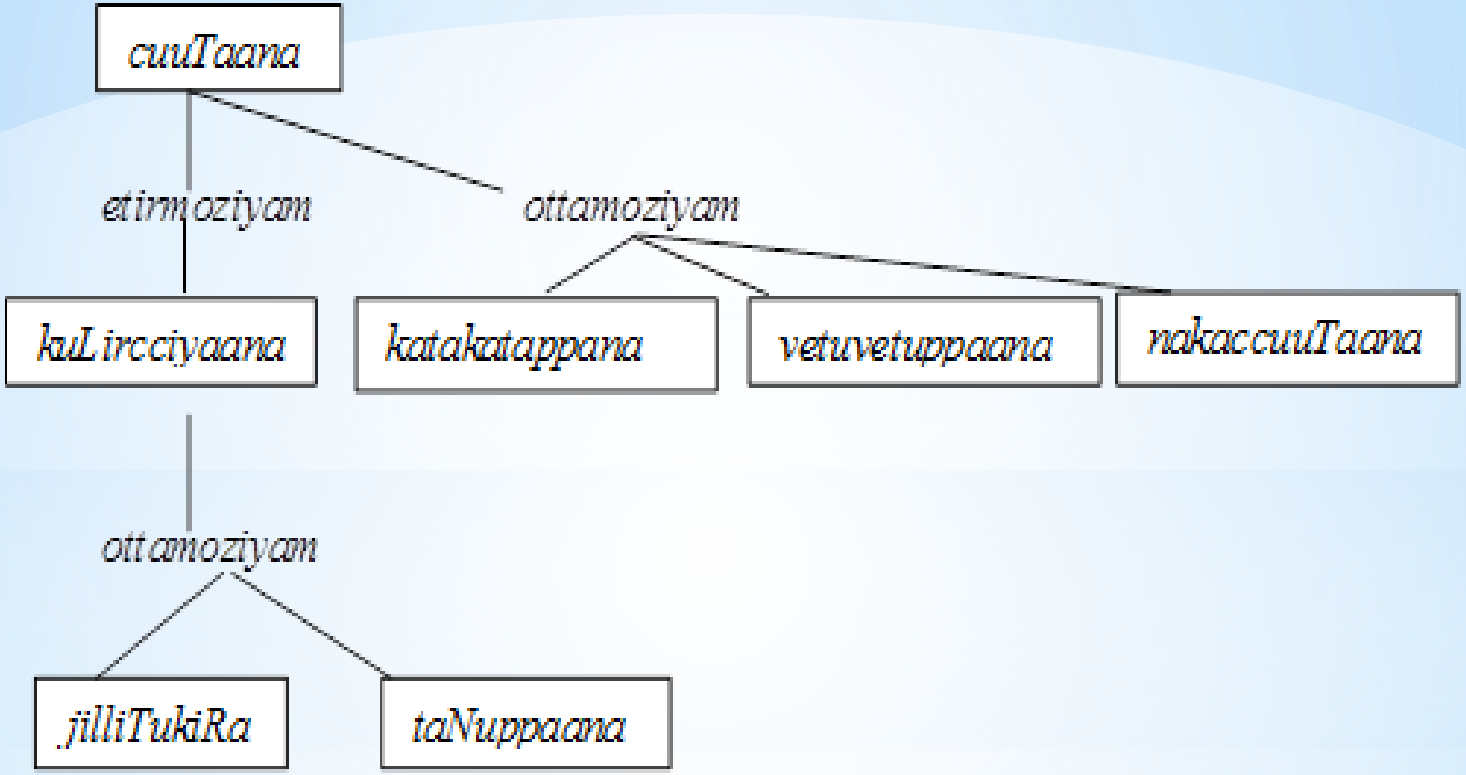
Relations	POS linked	Example
Antonymy (gradable i.e. contrary)	Adjective-adjective	*azakaana* 'beautiful': *kuruuramaana* 'ugly'
Antonymy (non-gradable i.e. contradictory)	Adjective-adjective	*uyiruLLa* 'alive': *cetta* 'dead'
Derivational	Adjective-noun	*azakaana* 'beautiful': *azaku* 'beauty'
Attributive	Noun-adjective	*vaTivam* 'size': *cinna* 'small'
Relational	Adjective-noun	*poruLaataara* 'economical': *poruLaataaram* 'economy'
Similarity	Adjective-adjective	*paaramaana* 'heavy': *kanamaana* 'heavy'
Derivational	Noun-adjective/ adverb	*azaku* 'beauty': *azakaana* 'beautifull', *azakaaka* 'beautifully'
Similarity	Adverb-adverb	*veekamaaka* 'fast' : *viraivaaka* 'fast'

* Organization of Abstracts in VOT



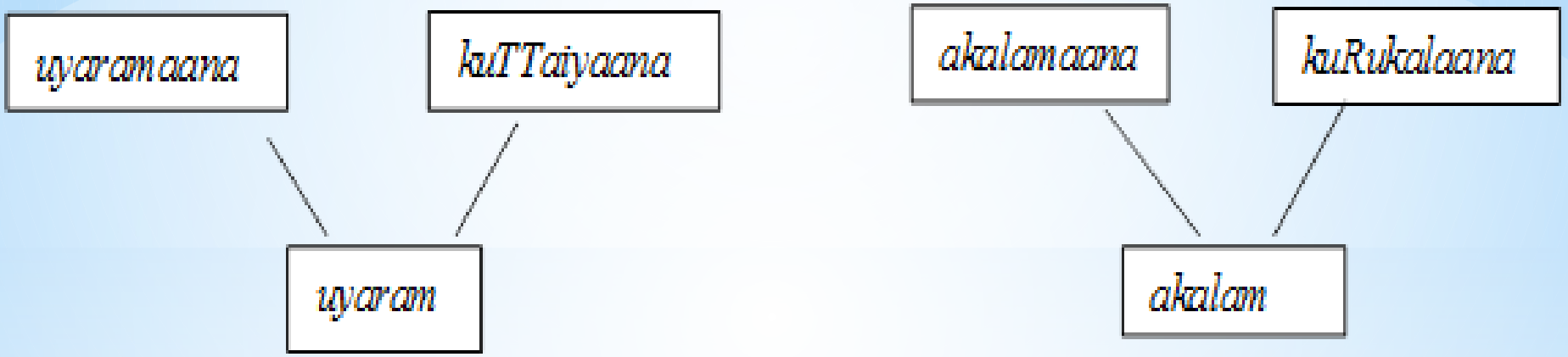
Antonymy in Adjectives

* Organization of Abstracts in VOT



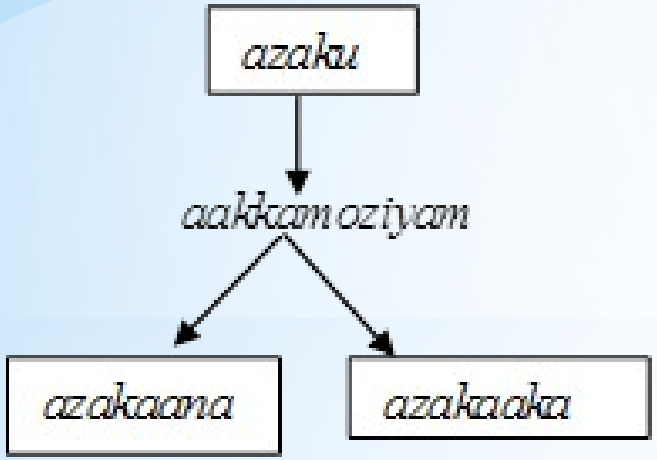
Similarity in Adjective

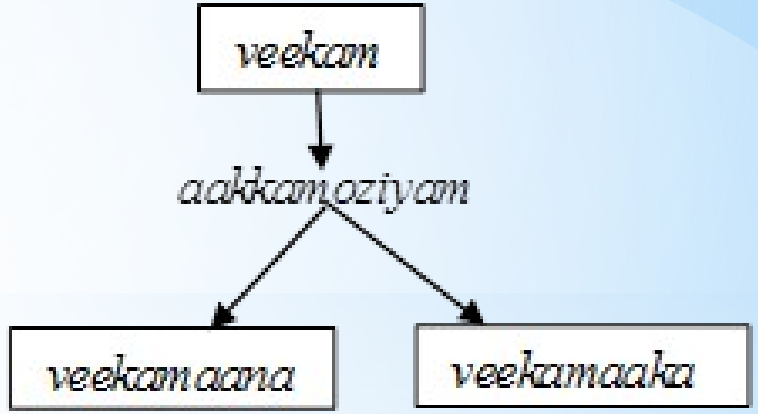
* Organization of Abstracts in VOT



Markedness in Adjective

* Organization of Abstracts in VOT



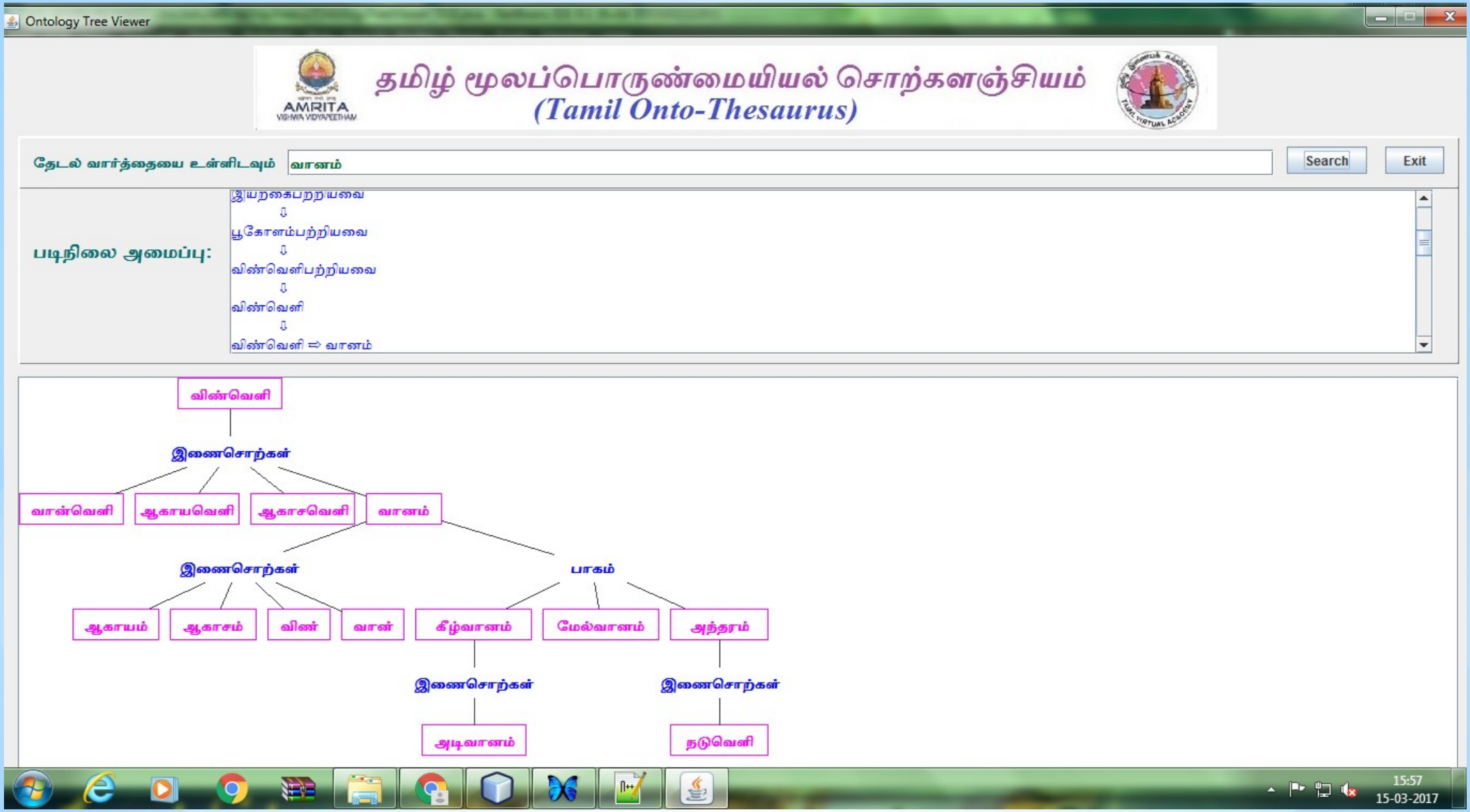


Derivative relation of Adverbs noun

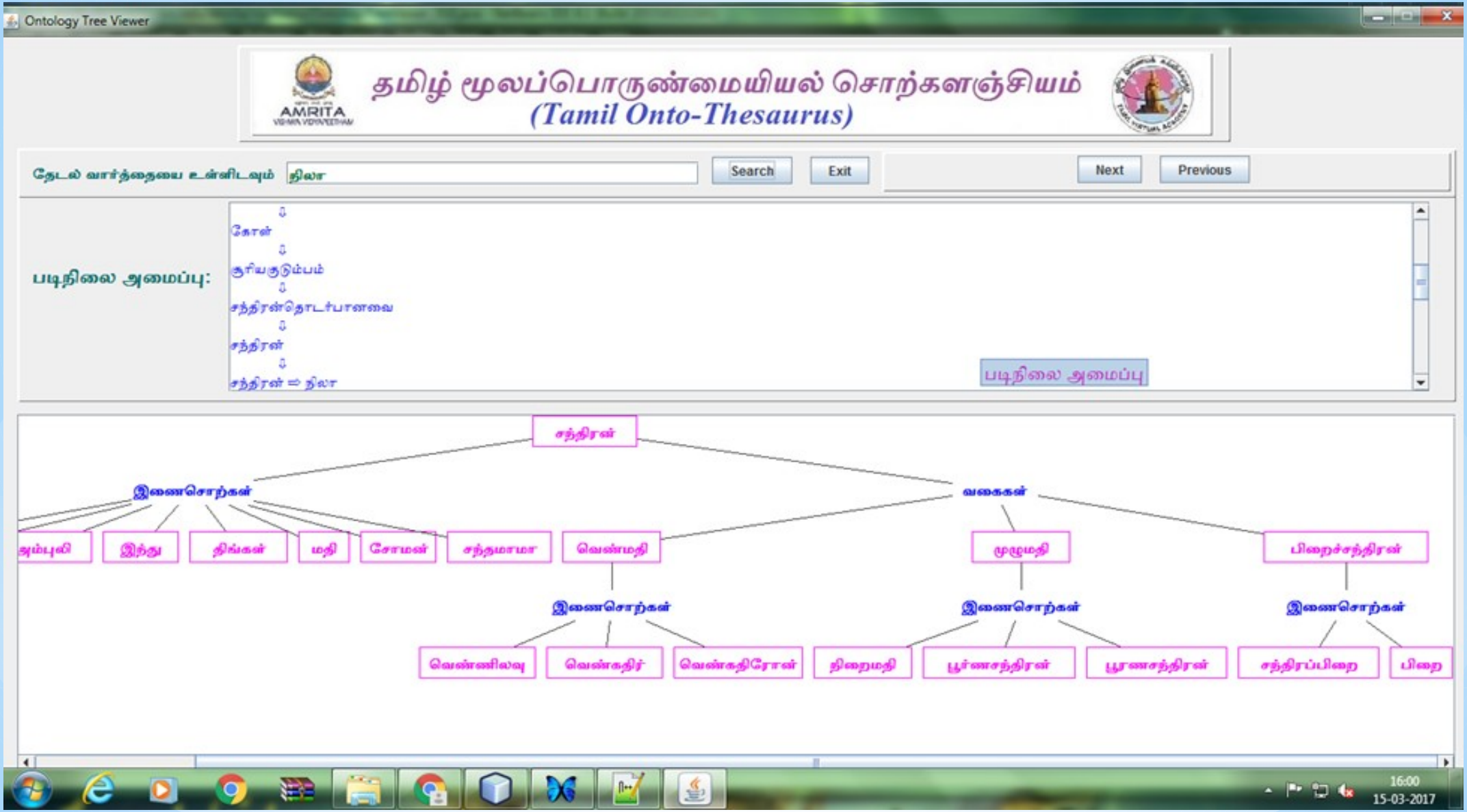
* Organization of Abstracts in VOT

*Nida listed relational concepts under the headings spatial, temporal, deictic, logical, etc. VOT has also followed his approach.

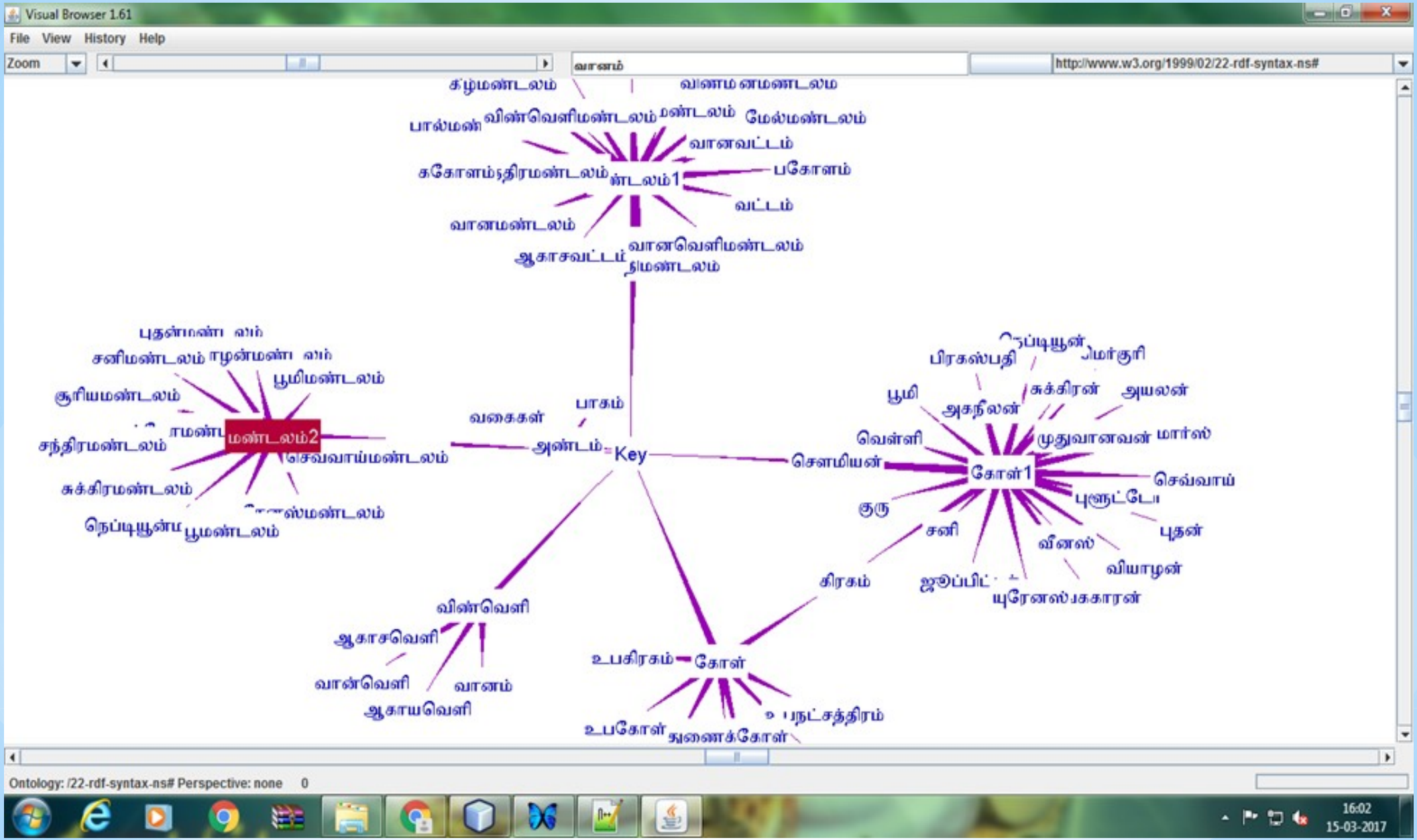
*** Organization of Relational in VOT**



Tree Viewer for VOT



Tree Viewer for VOT



Visual onto-thesaurus

THANKS TO THE ORGANIZERS AND LISTENERS